236330 – Introduction to Optimization and Deep Learning
HW1
Uri Kirstein – 311137095
Pavel Rastopchin- 321082026

Task 1:

Gradient calculation:

$$f_1(x) = \phi(Ax) \rightarrow u = Ax \rightarrow du = Adx \rightarrow df = \nabla\phi(u)^T du \rightarrow df = \nabla\phi(\text{Ax})^T Adx \rightarrow g^T = \nabla\phi(Ax)^T A$$
$$\rightarrow g_{f1} = A^T \nabla\phi(Ax)$$

Hessian calculation:

$$g_{f1} = A^T \nabla\phi(u) \rightarrow dg = A^T \nabla^2\phi(u)du \rightarrow dg = A^T \nabla^2\phi(Ax)Adx \rightarrow H_{f1} = A^T \nabla^2\phi(Ax)A$$

Task2:

Gradient calculation:

$$f_2 = h(\phi(x)) \rightarrow u = \phi(x) \rightarrow du = \nabla\phi(x)^T dx \rightarrow df = h'(u)du \rightarrow df = h'(\phi(x))\nabla\phi(x)^T dx \rightarrow g^T$$
$$= h'(\phi(x))\nabla\phi(x)^T \rightarrow g_{f2} = \nabla\phi(x)h'(\phi(x))$$

Hessian calculation:

$$dg = a'b + ab' = \nabla^2\phi(x)dx\, h'(\phi(x)) + \nabla\phi(x)h''(\phi(x))\nabla\phi(x)^T dx$$
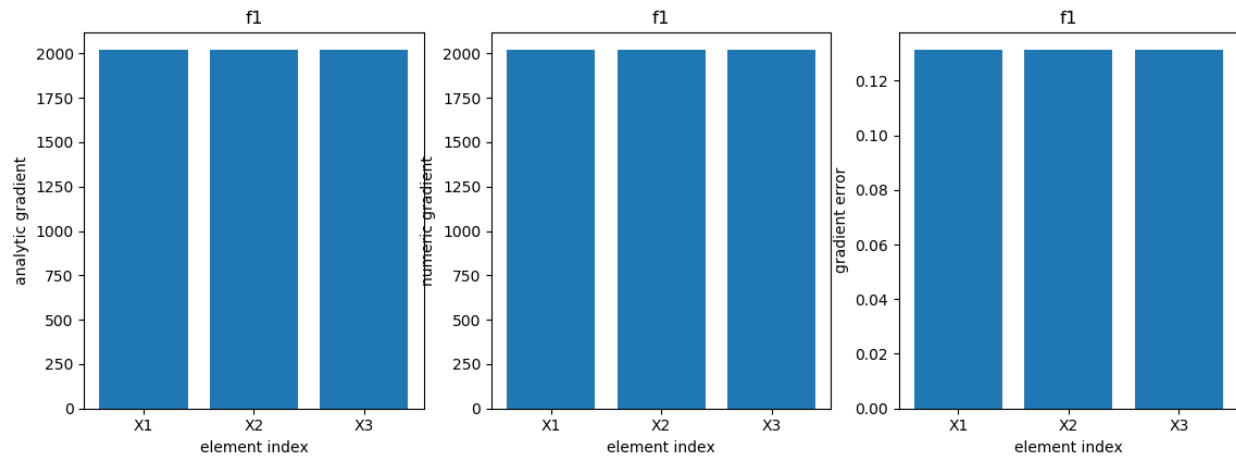$$\Rightarrow dg = \left(\nabla^2\phi(x)\, h'(\phi(x)) + \nabla\phi(x)h''(\phi(x))\nabla\phi(x)^T\right)dx$$
$$\Rightarrow H_{f2} = \left(\nabla^2\phi(x)\, h'(\phi(x)) + \nabla\phi(x)h''(\phi(x))\nabla\phi(x)^T\right)$$

Task 5:

Final results:

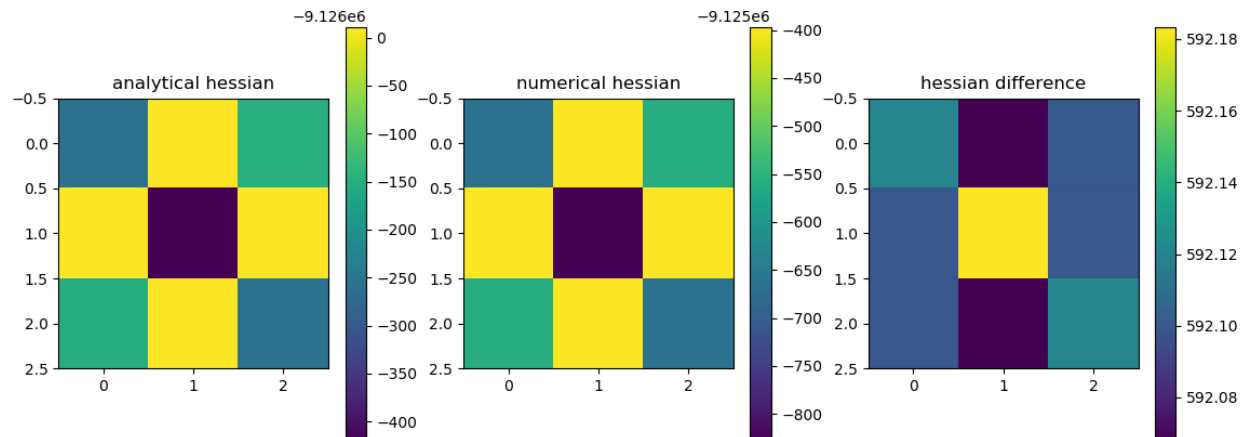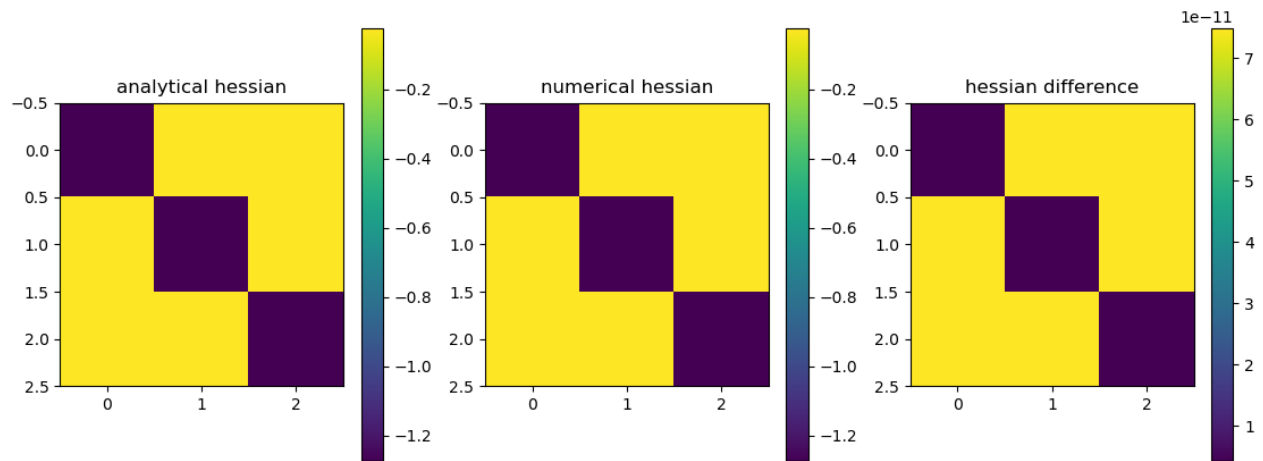|	Infinity norm	
	epsilon of minimal gradient error	epsilon of minimal hessian error
f1	4.24E-23	4.24E-23
f2	6.81E-18	3.04E-17

Plots:

f1 error of gradient numeric function by element index

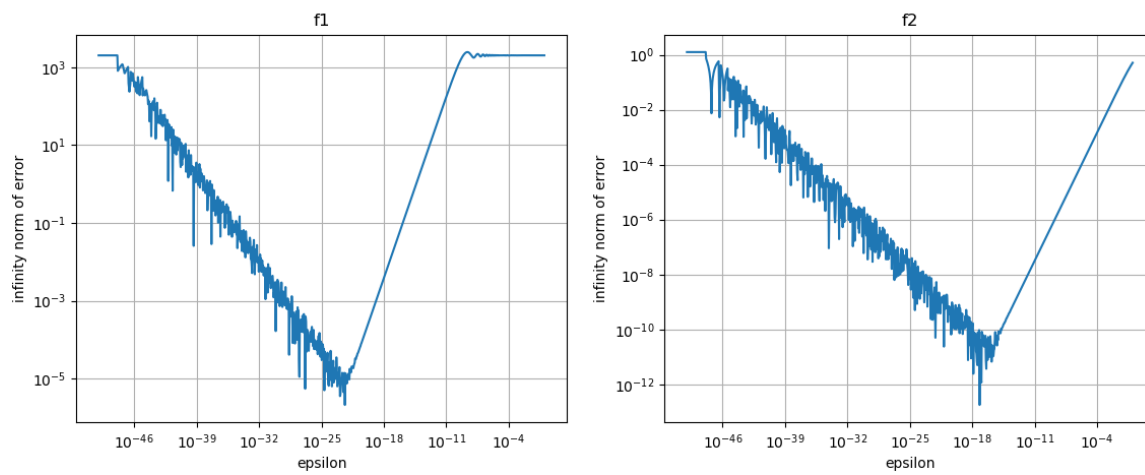## f2 error of gradient numeric function by element index
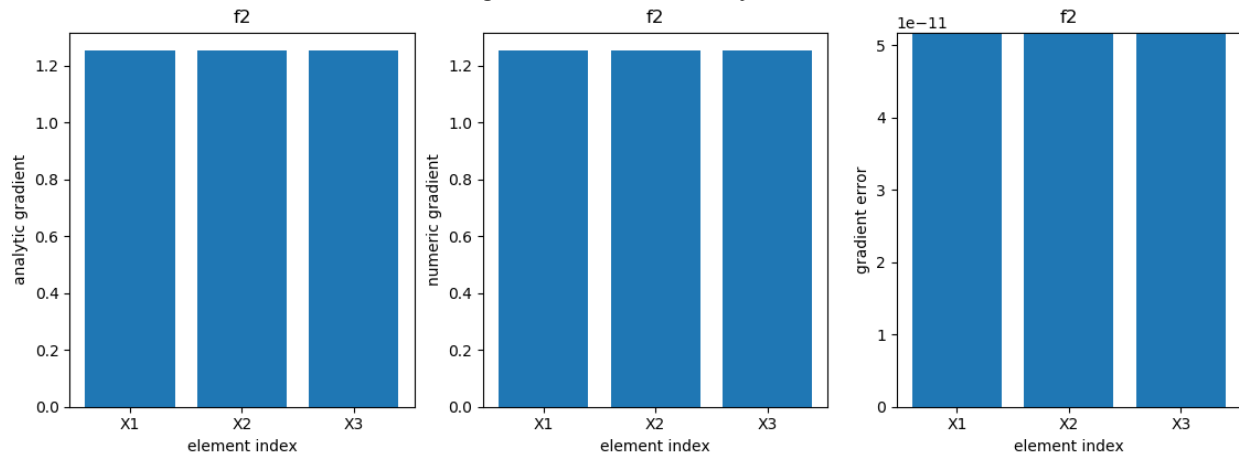
f1 hessians and their difference

f2 hessians and their difference